PCT

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification [6] : G01N 33/53	A1	(11) International Publication Number: **WO 98/12559**
		(43) International Publication Date: 26 March 1998 (26.03.98)

(21) International Application Number: PCT/US97/16738	(81) Designated States: AU, CA, JP, US, European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE).
(22) International Filing Date: 19 September 1997 (19.09.97)	
(30) Priority Data: 60/026,415 20 September 1996 (20.09.96) US	**Published** *With international search report.*

(60) Parent Application or Grant
  (63) Related by Continuation
    US   60/026,415 (CIP)
    Filed on 20 September 1996 (20.09.96)	

(71)(72) Applicant and Inventor: DEMERS, James, P. [US/US]; Apartment 6N, 380 Riverside Drive, New York, NY 10025 (US).	

(54) Title: SPATIALLY ADDRESSABLE COMBINATORIAL CHEMICAL ARRAYS IN CD-ROM FORMAT

(57) Abstract

A method for preparing very large spatially-addressable arrays of chemical compounds by light-directed synthesis is provided, wherein the light is provided by a laser and the compounds are arrayed on a rotating disc in a CD-ROM format. A method for assaying the resulting array with a CD-ROM mechanism is also provided.

Spatially Adressable Combinatorial Chemical Arrays
in CD-ROM Format

Field of the Invention

The invention has applications in the fields of chemistry, drug discovery, and genomics. The invention relates to: (1) methods for conducting chemical reactions on a support surface with spatial selectivity; (2) the spatially-addressable arrays of compounds (chemical libraries) produced thereby; and (3) methods for detecting specific members of these arrays of compounds, including methods of assaying for the specific binding of substances such as receptors, antibodies or other ligands, as well as methods of assaying for biological activity or other desirable property. The invention employs laser light under computer control for both the synthesis and assay of the arrays, using recordable compact-disc technology previously developed for optical storage of computer data.

Background of the Invention

The advantages of combinatorial chemistry for the rapid generation of chemical compounds for pharmaceutical screening are well established. The single greatest strength of combinatorial chemistry is that it makes possible the generation of libraries of huge numbers of compounds in a relatively short period of time.

Several methods have been employed for the preparation of such libraries, e.g., solution-phase synthesis; solid-phase synthesis on polymer beads or other divided supports; synthesis on soluble, precipitatable polymer supports; and synthesis on planar supports. All of these methods are capable of generating large libraries (i.e., libraries containing a large number of compounds), but none of them are amenable to rapid screening of the libraries for binding activity, biological activity, or other desirable properties.

Current methods for screening libraries include those in which individual library members, or small groups of members, are assayed in microtiter plates, e.g., by scr ening for a desired activity or for binding to a specific binding

- 2 -

partner, such as a receptor or antibody or other ligand, but the number of compounds that can be assayed at once is on the order of 10^2 to 10^4: one compound per well in a 96-well microtiter plate screens at most 96 compounds, while twenty compounds per well in a 384-well plate screens about 7680 compounds. The latter approach, while permitting higher throughput, requires secondary screening to identify the active species in any given well. For example, screening a library containing all of the 3.2 million possible pentapeptides which could be made from the twenty natural amino acids (i.e., a 3.2-million-member library) by these methods would require 500 to 3400 plates. Screening a library of the 64 million possible hexapeptides would require 10,000 to 68,000 plates. Robotic systems are available for microtiter plate assays, but screening a large library by such a method remains a massive undertaking. Furthermore, in a pharmaceutical discovery environment this represents only one of the many assays an organization might wish to conduct.

An alternative to testing individual compounds is the testing of complex mixtures, with various "deconvolution" strategies being employed to deduce the active species. These strategies have in common the re-synthesis and re-testing of successively less-complex mixtures. In addition to the great effort involved, the testing of complex mixtures is limited by the low concentration of any individual species in the mixture, and is susceptible to false positive results from the additive effects of large numbers of weakly active species. In practice, deconvolution has had limited success in probing large libraries of compounds. See L. Wilson-Lingaro, *J. Med. Chem.*, **39**, 2720-2726 (1996); and D. A. M. Konings, *J. Med. Chem.*, **39**, 2710-2719 (1996), and references therein, for a discussion of deconvolution strategies.

Ideally, the probing of a combinatorial library would be conducted in a single operation, with the active members of the library being in some way "pointed out" of the vast population of compounds by th assay. Two of th curr nt methods meet this requirement. Both methods employ solid-phase synthesis, and both require that the library members r main attached to the

- 3 -

solid support. In on m thod, a library of compounds bound to polystyren beads is prepared by the split-mix method. The library is assayed in a single batch, by being exposed to a molecule of interest, such as a receptor, enzyme, or other specific binding partner. Any beads to which the molecule binds are visualized (e.g., by a colorimetric assay), and beads so identified are selected and the structure of the library member attached to the bead is determined. This can be done by sequencing if the compound is a peptide or nucleic acid, as described, e.g., in U.S. patent 5,382,513 (incorporated herein by reference). The structure of the compound on the bead may in some cases be deduced from spectroscopic evidence (see, e.g., U.S. Patent 5,382,513), or by decoding a chemical tag that reveals the chemical history of the bead, as described in patent application WO 95/24186 (incorporated herein by reference). The method is in principle capable of screening very large libraries, limited only by the number of beads one is willing to examine. In practice, libraries of 10^4 to 10^6 members can be dealt with in this fashion.

The second approach involves physically locating a compound or compounds in a spatially addressable array of compounds on a planar support. In this approach, a compound's identity is revealed by its location in the array. One method of this type employs an array of compounds generated by light-directed synthesis, as first disclosed by Fodor et al. in Science, 251, 767-773 (1991), in which a fraction of sites on a planar support carrying photo-detachable protecting groups is exposed to light through a photolithographic mask, and the fraction of sites thus deprotected are functionalized with a specific monomer or building block, itself carrying a photo-detachable protecting group. The process is repeated with the mask in a different position or orientation, or with a different mask, and a second monomer or building block is attached to the support and/or to the first monomer residues. After numerous such cycles, with careful attention to the pattern of masking, an array of compounds is built up on the support. The final array is completely deprotected, and xposed to the ligand of int rest. Binding of th ligand is visualized by immunofluorescence, using antibodi s against the ligand which

- 4 -

are tagged with a fluorescent dye. Under a fluor_scence microscope, any location in the array to which the ligand has bound is visible as a fluorescent area, and the x-y coordinates of the area reveals the identity of the library member to which the ligand was bound. This technology, as applied to polypeptide and oligonuceotide synthesis, is known as Very Large Scale Immobilized Polymer Synthesis, or VLSIPS. It is described in US Patents 5,143,854, 5,413,939, 5,424,186, and 5,527,681, all of which are incorporated herein by reference.

The photolithographic method of synthesis, however, is cumbersome, and requires a substantial investment in very specialized equipment. A further investment in a fluorescence microscope or a specialized scanner is required for the assay, and highly skilled technicians are required, at least to conduct the synthesis aspect of the process. The scale of library synthesis is limited by the size of the masks and by the translational reach of the scanning device, which together limit the accessible surface for synthesis to a few square centimeters. In practice this technique is presently limited to arrays of 10^4 to 10^5 compounds. For these reasons the method is not routinely employed; see G. Jung and A. G. Beck Sickinger in *Angewandte Chemie*, **31**, 367 (1992).

As an alternative to photolithography, the use of directed laser light to conduct light-directed synthesis has been described in US Patents 4,719,615 and 5,318,679 (both of which are incorporated herein by reference). In the latter patent, a rectangular array support is either held stationary or translated, while a laser beam is scanned across the array by means of a rotating mirror, in the manner of an ordinary laser printer. This provides an alternative to the photolithographic masking approach, but fluorescence microscopy is still relied upon as an assay method.

Another alternative method for the synthesis of spatially addressable arrays utilizes ink-jet printing technology to spray micro-droplets of reagent solutions onto a substrate. This m thod is disclosed in US Patents 5,474,796 and 5,449,754, both of which are incorporated her in by r fer nce. This method is in theory capabl of preparing arrays of 10^7 compounds, but a m thod of

indexing such large arrays is not disclosed. Again, fluorescenc microscopy is the preferred means of conducting an assay.

There remains a need for reliable methods of generating very large, very high-density arrays of chemical compounds, on the order of about 10^8 or more compounds, along with a method of rapidly screening such arrays for chemical properties of interest, such as binding to antibodies, cellular receptors or other ligands, catalytic activity, or inhibition of enzymes.

Brief description of the invention

The present invention provides an alternative method of preparing and assaying a spatially addressable array of compounds on a planar support. By the method of the present invention, light-directed synthesis is carried out using laser light, but with the advantage that the synthesis support is in the form of a spinning disc. The ability of modern mechanical-optical systems to direct laser light to very small areas with very high precision makes it possible to selectively and reproducibly irradiate over 10^9 individual locations on a single 5.5-inch disc, as is done for example when reading a CD-ROM. Thus, even with ten-fold redundancy to allow for statistical errors and noise, light-directed synthesis with laser light in the CD-ROM format might enable the preparation of arrays (referred to herein as CD-arrays) of over 10^8 compounds on a single support disc.

In the present invention, a disc-shaped CD-array, functionalized on its surface with reactive groups which are blocked by laser-removable protecting groups, is rotated under a laser beam which is scanned radially across the disc. The radial position of the laser beam, the angular position of the disc, and the on/off status of the laser are all under the control of a computer, which "writes" to the disc in the usual fashion. Wherever the CD-array is "written to", *i.e.* irradiated, the protecting groups are removed, activating "synthesis sites" where chemical reactions can be carried out. These synthesis sites, arranged in annular or spiral tracks, correspond to the "pits" encoding digital information on a conventional CD-ROM. Recordable CD-ROM drives capable of carrying out most of th mechanical aspects of the present inv ntion are commercially

available.

The term "CD-array" as used herein should be understood to refer to any circular disc bearing synthesis sites or synthetic molecules arrayed in concentric or spiral tracks. The term is not intended to be limited to discs having the dimensions, track spacing, formatting, and feature sizes of commercial CD-ROM discs. The commercial CD-ROM format represents the presently known limit to the resolution of the CD-arrays of the present invention. The analogy to CD-ROM format is made for convenience, and to illustrate the magnitude of the advantage of the present invention over the current VLSIPS method. CD-arrays having synthesis sites on the order of ten to twenty microns in diameter (100 to 400 square microns in area), for example, are intended to be within the scope of the present invention. Such "low-density" embodiments of the invention will not be capable of displaying 10^9 compounds, but they will require less precise and less sensitive equipment for their use. Less dense arrays will also be amenable to less complex formatting schemes, such as the use of polar coordinates in analogy to the X-Y coordinate system used in VLSIPS. The disc diameter may also vary widely, but is preferably from about one inch to about 12 inches.

By activating selected synthesis sites, and subsequently exposing the disc to a chemical monomer reagent A' capable of reacting with the activated synthesis sites, a chemical fragment A is attached to the selected sites. This fragment will usually carry a functional group, which will usually be masked by a laser-removable protecting group. The process of irradiating selected synthesis sites is repeated, and the disc is exposed to a second monomer reagent B', capable of reacting with activated synthesis sites to attach fragment B. Where these sites were previously selected for attachment of fragment A, reagent B' will covalently attach fragment B to the functional group of A. The cycles of irradiation and reaction are repeated, with the computer that controls the laser and disc rotation also maintaining a databas of the irradiation and chemical exposure history of ach synthesis sit , until the desired library of compounds is ass mbled from the monomer r agents. The entire disc, or only

- 7 -

selected synthesis sites, may then be irradiated to remove any remaining protecting groups

The disc is then assayed to detect library members having some desirable property, such as the ability to bind selectively to an analyte. Bound analyte is optionally made more detectable to laser light with a visualization reagent or process, and synthesis sites carrying bound analyte are then detected by "reading" the disc with a laser, in a manner analogous to the reading of a CD-ROM.

The present invention is particularly suited to sequencing DNA by hybridization to arrays of oligonucleotides, as described for instance in U.S. Patent 5,525,464, the contents of which are hereby incorporated by reference. In this embodiment of the present invention, fluorescently or otherwise labeled DNA fragments are allowed to hybridize to an array of oligonucleotide probes on a disc, and by identifying the probes to which the DNA has hybridized the sequence of the DNA may be deduced. As described in US 5,525,464, one would ideally carry out hybridizations to all possible 11-mer oligonucleotide probes, but the technical difficulty of preparing such an array is acknowledged to be a limitation. The present invention removes this limitation, and makes possible for the first time the preparation of an array of all 4,194,304 11-mer oligonucleotides. Indeed, the present invention makes possible an array of all 16,777,216 possible 12-mers, which would reduce the number of target DNA fragments needed for sequencing a given gene or genome. Arrays of even longer oligonucleotides are possible (there are just over 10^9 possible 15-mers), but end mismatches become difficult to distinguish as the oligonucleotide length increases.

Detailed description of the invention

The array disc comprises a first layer, hereinafter referred to as the synthesis layer, and preferably incorporates a second layer, hereinafter referred to as the reflective layer, located below the synthesis layer. At least one additional layer will be present, for purposes of mechanical strength, for protection of a reflective layer or of the synthesis layer, and to constrain liquid

- 8 -

reagents to the synthesis layer. The synthesis layer is initially functionalized with reactive groups, such as amino groups, which are blocked by a laser-removable protecting group. The materials of which the array disc is constructed are not critical, they may for example be metallic, polymeric, ceramic, vitreous, or composites thereof. The only requirements of the materials are that, where they are exposed to solvents and reagents during the synthesis of the library members, they be effectively insoluble in and inert to those solvents and reagents. The materials must also render the disc sufficiently rigid to avoid deformations that would prevent accurate irradiation of the synthesis sites. Where a reflective surface is separated from the synthesis surface by a layer of material, or where a transparent disc is employed, this material must also be transparent to the frequency of the laser used to assay the array, referred to herein as the interrogating laser.

Methods of preparing functionalized surfaces, including functionalized glass, metal, and polymer surfaces, are well known in the art. The functional groups may be attached to the surface by a spacer of any desirable length. The spacer may possess certain desirable properties, such as low non-specific binding to proteins or nucleic acids, as appropriate for the particular assay being contemplated by the practitioner. The spacer may be linear or branched, and may optionally carry a plurality of functional groups. The patents incorporated herein by reference disclose or reference exemplary methods to those skilled in the art. The fluorocarbon surface masking method disclosed in US patent 5,474,796 may optionally be employed, wherein an initial laser ablation of a fluorocarbon layer is used to constrain the synthesis sites to pre-defined locations.

It will be appreciated by those skilled in the art that the disc will require formatting in some manner, in order to define tracks, sectors, frames, addresses, and other features. These features provide tracking error detection, frame sync signals, data addresses, and other information to the read-write device that enable accurate and reproducible writing and reading of data in a CD-ROM, and similar information will be required for reproducible and accurat

- 9 -

irradiation and subsequent assaying of the CD-array. A variety of CD data formats are available, and the present invention contemplates the use of any of these formats, or other methods of locating a selected synthesis site.

Certain formatting features can be incorporated into the structure of the disc itself, for example by pressing, etching, or laser ablation, as is well-known in the art. These features are usually detectable by the writing and reading laser optics, but may be read and utilized by other components of the read-write device, such as a separate diode laser dedicated to tracking. The present invention anticipates that the disc may be spun with either constant linear velocity or constant angular velocity, at the discretion of the practitioner.

Other formatting features can be written by the "writing" laser in a formatting step prior to the initiation of library synthesis, in a manner analogous to the formatting of a hard disc drive. Synthesis sites activated in the formatting step will then be reacted with a "formatting reagent", in order to attach a formatting moiety that provides a detectable signal to the interrogating laser. The formatting moiety will be chosen to be resistant to removal by the chemical solvents and reagents subsequently used by the practitioner in the synthesis of the library. Detectability might be inherent to the formatting reagent, or might be introduced in subsequent operations that modify the formatting moiety. As examples of inherent detectability, if the library is to be assayed by attenuated reflectance, the formatting reagent may comprise one or more unreactive dye molecules absorbing the frequency of the interrogating laser. If the library is to be assayed by interferometry, the formatting reagent may comprise an unreactive polymer of sufficient size to provide a detectable change in phase of the interrogating laser light. As an example of introduced detectability, the formatting steps could comprise introducing biotin as the formatting moiety, followed by avidin binding to the biotin and silver staining of the avidin to opacify the biotinylated synthesis sites.

The reading of a CD-array for detection of binding of analyte will in some cases involve scanning a mostly "blank" disc for th occasional signal. It will b appreciated by those skilled in the art that the "eight to fourteen modulation"

- 10 -

or EFM encoding characteristic of current audio and CD-ROM disc technology is not required in such cases, since binding of analyte to adjacent synthesis sites is unlikely. Thus, the present invention anticipates that the computer may have direct control of the writing mechanism during synthesis, and also may have direct access to the primary signal generated by the reading optics during the assay, bypassing much of the electronic circuitry present in conventional CD-ROM devices.

The term "light-directed synthesis" as used herein refers not only to photochemistry, but also to thermal reactions that are induced by laser irradiation, and catalyzed reactions where the catalyst is photo-generated. Reactions that either deprotect or create reactive functional groups are contemplated to be within this definition.

The deprotection (activation) of the synthesis layer and of the growing library members by presently available photochemical (as opposed to thermal) means would be impractical using the diode lasers available at the present time, since diode lasers are not currently capable of generating ultraviolet light of sufficient intensity to carry out photochemistry at an acceptable rate. Ultraviolet laser diodes of sufficient intensity would be ideal for the present application. Photochemical deprotection, therefore, is preferably accomplished at the present time by replacing the diode laser found in a commercial recordable CD-ROM device with a more powerful "bench top" laser, and delivering the laser beam to the disc write head by reflection or by means of fiber optics. US Patent 5,318,679 describes appropriate laser equipment for conducting photochemical deprotection for purposes of light-directed synthesis, and similar equipment is widely available from commercial sources. The laser will be chosen to provide a frequency of light appropriate for the particular photo-removable protecting groups being employed. With the laser being external to the CD-ROM device, the practitioner has the option of switching lasers, making it possible to employ a plurality of protecting groups. It will be und rstood that wh n reference is mad herein to the "radial motion" of the laser, the expression refers to motion of th read-write head, which may or may

- 11 -

not actually comprise a laser. The read-write head may comprise a lens at the end of an optical path, with the laser itself remaining stationary.

The size of the synthesis area will in most cases be determined by the degree to which the laser light is focused, which is dependent on the focal length of the lens through which the light is delivered and on the distance from the lens to the synthesis layer. The size of the synthesis sites can be varied to suit the resolution of the overall system and the number of library members to be prepared on the disc.

The intensity and duration of a laser pulse required for quantitative or near-quantitative deprotection or activation will be dependent on numerous factors: the power and frequency of the laser, the amount of power actually delivered to the synthesis sites by the system optics, the size of the synthesis site, and the photochemical properties of the protecting group chosen by the practitioner. Determination of the most effective power level and pulse length can be determined by known methods, for example those disclosed in U.S. patent 5,143,854, and is within the ability of one skilled in the relevant arts.

Alternatively, photo-removable protecting groups sensitive to the wavelengths available from current diode lasers would be advantageous. Such groups have not previously been reported, since there has been no need for them, but they could be designed by modification of the absorption maxima of known photoremovable protecting groups.

Some of the most effective currently employed photo-removable protecting groups, such as the nitroveratyloxycarbonyl (NVOC) group, require a carbonyl group scavenger to be present in solution for optimum yields. Photo-removable silyl protecting groups, while not requiring a scavenger, nonetheless function more efficiently in a polar solvent. See V. N. R. Pillai, *Synthesis*, 1, (1980), incorporated herein by reference, for a general discussion of photo-removable protecting groups. See U.S. patents 5, 486,633 and 5,489,678 (both incorporated herein by reference) for photo-removable protecting groups particularly useful for solid-phas synthesis of peptides and nucleic acids, and see U. Zehavi, *Adv. Carbohydrate Chem. and Biochem.*, 46, 179-204 (1988)

- 12 -

for photo-removable protecting groups useful for carbohydrate synthesis.

Where such protecting groups and/or solvents are to be employed in the present invention, the CD-array structure will preferably incorporate an optically flat cover layer transparent to the wavelength of the activating laser, sealed at the circumference and at the inner annular edge, and will incorporate ports for introduction and withdrawal of reagent solutions. The synthesis layer, the seals, and the cover layer will define a disc-shaped space where reagent solutions may be maintained in contact with the synthesis layer during irradiation. An inner annular seal may optionally be a rotating seal, so that solutions might be pumped through the space under computer control without the need for removing the disc from the laser activation apparatus. The present invention also contemplates an alternative embodiment, in which a thin layer of solution or solvent is spin-cast onto the disk surface, using well-known technology developed for the preparation of thin films (See, e.g., C. W. Frank *et al*, *Science*, **273**, 912-915, and references therein). Non-volatile solvents will be preferred. It will be appreciated that such designs also permit the automated introduction of monomer reagents and solvents.

It will be appreciated that the synthesis layer need not be rigid, but may comprise a covalently attached fluid or liquid crystal phase, such as a polyethylene glycol, or alternatively may comprise a gel phase with solvent molecules incorporated therein. Such solution-like phases are well known in the art, and are frequently used where polymer beads are employed as the support in solid-phase synthesis. In certain cases, such a phase is expected to obviate the need for a solvent space or layer.

It will be appreciated that the presence of a cover layer and/or solvent will effectively alter the focal length of the objective lens of the CD laser, and that the refractive index and thickness of these additional materials will have to be precisely controlled and/or monitored, and compensated for by the CD mechanism, if maximum resolution is to be maintain d. Photo-removabl protecting groups that do not require a liquid phase would be preferred in the present invention. Protecting groups that photo-fragment to volatil byproducts,

for example, might be employed, with operations b ing conducted in a vacuum if required.

The deprotection of the functional groups may also be accomplished thermally by irradiation with conventional infrared lasers, for example the diode laser of a recordable CD-ROM device. Most rewritable CD-ROM devices rely on thermal effects to write data bits to the disc, using the localized heating induced by brief, highly focused irradiation. Momentary surface temperatures in excess of 300°C are obtainable with such devices. US patent 4,719,615 (incorporated herein by reference), together with the references therein, describes the technology and principles involved. Thermally removable protecting groups are known in the art; for example the t-butoxycarbonyl group (Boc group) decomposes to gaseous products at temperatures of about 180°C (J. S. Zambounis et al., Nature, 388, 131-132 (1997); V. H. Rawal and M. P. Cava, Tetrahedron Lett., 26, 6141 (1985); and H. H. Wasserman and G. D. Berger, Tetrahedron, 39, 2459 (1983)). It will be appreciated that where the thermal deprotection is not completed during a single irradiation period, multiple passes under the laser will eventually drive the reaction to the desired level of completeness, albeit at the cost of slowing down the overall process. Alternatively, the period of irradiation may be extended, however this will cause a larger area of the synthesis layer to be heated and deprotected, thereby reducing the achievable density of synthesis sites on the array. Partial deprotection at the margins of the synthesis site is to be expected, resulting in the synthesis of mixtures of compounds in the margin as the cycles accumulate. If the interrogating laser can selectively irradiate the central portions of the synthesis site, where deprotection is reliably complete at each cycle, this effect can be ignored. Alternatively, masking of the area around the synthesis sites by means of a fluorocarbon layer as described in US patent 5,474,796 can eliminate marginally heated areas. It is anticipated that thermal reactions other than deprotection reactions will be inducibl by the activating laser, for example acyl azides could be converted selectively to isocyanates as one step in a light-directed synthesis.

- 14 -

To initiate library synthesis, a pre-determined set of sites on a formatted CD-array disc is deprotected or otherwise activated by laser irradiation, and the array is then exposed to chemical reagents which attach the desired chemical moieties to the irradiated sites. The array is subjected to as many cycles of irradiation and chemical treatment as are desired. The computer may be used to schedule the processes, and is used to control the irradiation patterns and to store a record of the irradiation patterns and the chemical treatments that took place after each irradiation session. This record, or chronology, can be referred to later to enable the practitioner to deduce the structure of the library member at any given synthesis site. Alternatively, the entire chronology may be pre-determined by the practitioner, and employed as an instruction set by the computer.

As used herein, the term "spatial location" in reference to the location of a particular synthesis site or library member on the disc is equivalent to the "logical address" of the synthesis site, for purposes of associating bound analyte with the identity of the library member to which it has bound. The logical address is a locator which indicates a location on the disk with reference to formatting features, such as tracks, sectors, blocks, and the like. It will be appreciated that one need not know the actual physical locations of the library members on the CD-array disc in order to practice the present invention, unlike the prior art spatially addressable arrays that rely on x-y coordinates to encode compound identity.

The writing of a computer program capable of carrying out the tasks required by the present invention is within the ability of one skilled in the art of computer programming. The present invention contemplates that a variety of software designs will be applicable to the tasks of mapping the library onto the disc and directing the movements of the lasers, the CD-array disc, and optionally controlling reagent and solvent delivery to the CD-array.

After the desired number of irradiation and reaction cycles, the CD-array will usually b stripped of all remaining protecting groups. It will be appreciated that protecting groups not sensitive to removal by irradiation can be employed

- 15 -

at any chemical reaction stag , so as to mask reactive functional groups that are desired to be present at a later stage of synthesis. Such groups may, for example, be removable by chemical or thermal means. Protecting groups that are photoremovable, but not sensitive to the wavelength of light being employed for the rest of the library synthesis, can be similarly employed. Such protecting groups will be utilized, for example, where a branched oligomer is the desired product, and the branching point requires activation after one or more rounds of irradiation.

The CD-array is then exposed to a substance of interest, under conditions that permit selective binding to those library members that have the greatest affinity for the substance. (The substance of interest, usually a protein, nucleic acid, or other biomolecule, is hereinafter referred to as the analyte.) Methods for effecting the selective binding of an analyte are well-known in the art, and are routinely employed in analytical procedures and when supported compound libraries are probed for binding affinity. Likewise, methods for the hybridization of nucleic acids to immobilized libraries of oligonucleotides are well-known, for example see U.S. patent 5,002,867, incorporated herein by reference.

In the pharmaceutical field, the analyte will usually be a protein, for example a receptor, an ion channel, an enzyme, or a signal transduction protein. In the genomics field, or in diagnostics or forensics, the library will often consist of nucleic acid sequences, or sequences of nucleic acid analogues such as peptide nucleic acids, and the property of interest will usually be the ability of library members to hybridize to single-stranded DNA or RNA. The ability of such libraries to bind DNA- or RNA-binding proteins, such as transcription activators or repressors, will also be of interest in the biochemical and pharmaceutical fields.

Typically, the array is exposed to a solution of the analyte of interest, under appropriate conditions to permit any selective binding or hybridization to occur. Unbound analyte is rinsed away, and any bound analyte is then optionally rendered more detectable to laser light with a visualization reagent or process as described below.

- 16 -

Finally, the present invention provides for the detection of analyte bound to individual library members of the CD-array. The CD-array is "read" in the same way that a conventional CD-ROM disc is read. See L. Boden, *Mastering CD-ROM Technology*, John Wiley & Sons (1995), ISBN 0-471-12174-6, (incorporated herein by reference) for an overview of the relevant techniques of encoding and decoding information in the CD format. Light from the interrogating laser passes through a synthesis site, is reflected from the reflective layer below the synthesis site, and is detected after reflection back through the synthesis site. In an alternative embodiment, there is no reflective layer and the laser light is detected after passing through the disc. The reading may be accomplished by detecting changes in some property of the laser light which has passed through the synthesis sites, said change being induced by the presence of the analyte or by the presence of a visualization reagent bound to or associated with the analyte. This change in property is detected by an appropriate detector, and converted to a digital signal, usually a binary signal, which is recorded by the computer. The property of the light altered by the analyte may be intensity, as is the case in presently available CD-ROM readers, but may be some other property such as polarization angle, wavelength, or phase.

In an alternative embodiment of the present invention, fluorescence of the analyte or of a fluorescent label bound to the analyte is induced by the laser and is detected by an appropriate device, for example a charge-coupled device (CCD) or a photocell. An example of the detection with laser light of fluorescently labeled analytes bound to a spatially adressable array is the "GeneArray" scanner, manufactured by Hewlett Packard (Palo Alto, CA). This device analyzes a centimeter-size, rectangular surface carrying an array of fluorescently labeled analyte DNA bound to an array of oligonucleotides. The array is held stationary and the laser beam is scanned across the surface. The present invention, by employing a rotating disc format rath r than a stationary surface for the array of synthesis sites, makes it posible to carry out a similar operation over a far larger surface area. The "GeneArray" scanner is presently

- 17 -

priced at well over $100,000, whereas first g n_ration CD-ROM readers may presently be obtained for under $100.

In yet another embodiment, the interrogating laser beam may be linearly polarized, and the fluorescence polarization from a fluorescent label may then be detected. See M. E. Jolley, *J. Biomol. Screening*, 1, 33-38 (1996), and references therein, for a brief introduction to fluorescence polarization.

In one embodiment of the invention, the intensity of the light reflected from a synthesis site is noted by the computer. As the disc rotates, another synthesis site is brought into the laser beam, and the process is repeated. At any site where the analyte has bound to a library member, and where the site has been opacified to the laser beam by any of the methods described herein, the reflected signal is attenuated to a measurable degree. Scanning of the CD-array by laser reflectance therefore generates a signal whose intensity varies with the degree of opacification of each synthesis site. This signal is converted into a digital signal, which is processed by the computer. If the signal-to-noise ratio is sufficiently large, information about the relative degree of opacity, and hence about the relative binding affinity of the library member for the analyte, may be derived from the signal. A lower signal-to-noise ratio, on the other hand, might permit only a binary information signal (analyte bound/not bound) to be derived. The computer has access to the "chronology" of the array, which is a database containing the irradiative and chemical reaction history of each synthesis site in the array, and which defines the structure of the library member at each site. The computer can report this data for each synthesis site that has been rendered opaque, thereby correlating the structure with the binding of the analyte.

For detection based on such attenuation of the interrogating laser light, it will usually be necessary to modify the analyte so as to render it opaque to the laser beam. This may be accomplished by covalently attaching dye molecules to the analyte, such dye molecules being chosen to have a high absorption coefficient at the frequency of the laser. Preferably, larg numbers of dye molecules are attached, for exampl by way of branched or d ndrimeric linkers,

- 18 -

so as to amplify the absorbance.

Alternatively, a visualizing or "opacifying" protocol might be employed wherein a reporter molecule with affinity for the analyte, for example an analyte-specific antibody, is allowed to bind to the analyte bound to the CD-array. Those skilled in the art will appreciate that there exists a wide variety of well-known methods for visualization of bound analytes with reporter molecules, for example antibody-based methods such as ELISA, various "sandwich" assays, and the use of biotin and avidin conjugates. The method chosen, whatever the details, will result in any synthesis sites that carry the analyte being rendered relatively opaque to the interrogating laser beam. For example, dye molecules that absorb the laser light might be attached to the reporting molecule, or to a visualization reagent that binds to the reporting molecule. Alternatively, an insoluble dye might be generated by an enzyme that is bound to the analyte, and the precipitated dye relied upon to attenuate the interrogating laser beam. Other approaches might employ colloidal gold particles, or a precipitated metal such as silver, to disperse the interrogating laser beam. An example of the use of metal particles can be found in *Angew. Chem. Int. Ed. Engl.*, **36**, 1080 (1997), where nanoparticles of gold, platinum, and cadmium-selenium are employed to visualize photochemically deprotected synthesis sites with micron resolution. Colloidal particles may disperse or absorb light in a wavelength-dependent fashion.

These and other methods of selective opacification of sites carrying the analyte will be apparent to those skilled in the art. For an example of the use of attenuated reflectance to detect biomolecule binding, see Mandenius et al, *Anal. Biochem.*, **157**, 283-288 (1986). The combination of reporter molecules, plus any attached or bound visualization reagents that render the analyte detectable, is collectively referred to herein as a label, and the overall process of attaching the label to the analyte is referred to as labeling.

Laser light passing through a lay r of bound analyte will experience a change in phase, due to passag through th higher refractive index medium of the analyte, relative to light that has not passed through a layer of analyte.

- 19 -

Such a change of phas can be detected by interferom try. For an xample of interferometric detection of surface-bound biomolecules on a transmissive substrate, see U.S. patent 5,413,939, incorporated herein by reference. Suitable analyte labels for interferometric detection would present a long pathlength to the interrogating laser and/or a high refractive index. Suitable labels would, for example, comprise high-molecular weight polymer chains and high-refractive index materials such as halocarbons. Current CD-ROM and audio CD players, in fact, rely on interference effects to distinguish pits and lands on CD disks. For more recent examples of the reading of data on a reflective CD-ROM by means of interferometry, see U.S. patent 5,471,455, incorporated herein by reference. Although the present invention contemplates the use of a transmissive disc, a reflective disc is preferred because it enables the practitioner to more easily modify existing CD-ROM equipment for use in the present invention. Also, the reflective mode enhances signal strength because it entails passing the interrogating laser light twice through the synthesis sites, and through any detectable analyte attached thereto.

As described in the above-referenced patents, interferometry requires that the interrogating beam be split into a sample and a reference beam, and that the reference beam travel a path as nearly identical as possible to the sample beam, so that any phase change can be reliably ascribed to the presence of analyte. Generally, this requires that the reference beam be reflected from a location as physically close as possible to the synthesis site, to eliminate errors due to variations in the physical media. (Present audio and CD-ROM readers irradiate an area larger than the pit size, and use light reflected from the disc surface to either side of the information-containing track as the reference beam.) The formatting and/or track spacing of a CD-array disc intended for interferometric assay will thus have to allow adequate space near each synthesis site for the reference beam. It would also be advantageous to replicate every library member in such a way that repeating signals are generated from the spinning disc with a characteristic frequency, which facilitates the filtering of noise from the signal, as discussed in U.S. patent

- 20 -

5,413,939 referenced abov .

Although the analytes of interest are likely to be optically active, and thus capable of inducing rotation in a reflected beam of polarized light, the amount of rotation induced directly by these analytes is expected to be insufficient for detection in most cases.   It is anticipated, however, that "super-rotating" moieties, such as optically active helicenes, would provide effective visualizing reagents if attached at high enough density to the analyte.   Alternatively, the very presence of a layer of analyte alters the polarization state of reflected light, and thus ellipsometry might be relied upon to detect the presence of analyte, as disclosed by C. F. Mandenius *et al*, *Anal. Biochem.*, **137**, 106-114 (1984) and references therein.   Data detection by measurement of small changes in the rotation of polarized light is currently employed in some commercial recordable CD-ROM systems.   Representative examples may be found in U.S. patent 5,457,582 (incorporated herein by reference) and references therein.

U.S. patent 5,453,969 (incorporated herein by reference) discloses the use of conoscopic holography to read data from an optical disc, wherein the angular spectrum of reflected polarized light contains the desired information. A birefringent crystal rotates the polarization of the reflected light as a function of the angular spectrum, and a polarizer/analyzer then projects an image of the beam which exhibits fringes.   The spacing of the fringes is a function of the angular spectrum, hence detection of the fringe pattern constitutes recovery of the original data.   Given that a layer of attached analyte will alter the angular spectrum of reflected polarized light (Mandenius et al, *Anal. Biochem.*, **157**, 283-288 (1986), referenced above, employed polarized light in their reflectance assay), a detector employing conoscopic holography should also provide a suitable means of assaying the CD-array.

It will be appreciated that the method of assay provided by the present invention does not require that the CD-array be prepared by laser-directed synthesis. For example, mass production of multiple copies of a CD-array, as might be desirable with a CD-array of bound oligonucleic acids intended for

- 21 -

sequencing by hybridization, might be more economically achieved by another method, such as photolithography or microprinting.

It is intended that all patents and publications referred to in this disclosure are incorporated by reference, whether or not explicitly so indicated in the text.

Summary

The present invention is a novel approach to the preparation and screening of large compound arrays. There are two notable features of the present invention which provide significant advantages over prior art methods:

(1) It relies largely on existing chemical methods, such as light-directed synthesis of compound arrays, and may utilize existing technology, such as commercial CD-ROM hardware, with only a moderate degree of modification. The present invention does not require the practitioner to invest in expensive equipment.

(2) The novel employment of a spinning disc for the substrate, in place of the linearly translated or scanned rectangular substrates of the prior art, greatly increases the surface area that can be accurately and reproducibly addressed with a laser. This makes possible an order-of-magnitude increase in the number of compounds that can be prepared and screened in a convenient and economic manner.

Industrial utility

The present invention may be employed for the synthesis and assay of oligonucleotide arrays, or arrays of oligonucleotide analogues such as peptide nucleic acids. These arrays have utility in diagnostics, forensics, and in gene mapping and sequencing. See, for example, E. Pennisi, Science, 272, 1736-1738 (1996); D. J. Lockhart et al., *Nature Biotechnology*, 14, 1675-1680 (1996); M. Chee *et al., Science*, 274, 610-614 (1996); and A. Goffeau, *Nature*, 385, 202-203. It may also be employed for the synthesis of peptides, arrays of which have utility for antibody profiling, diagnostics, and screening for pharmaceutical activity. Another utility of the present invention is the synthesis and screening of libraries of small (molecular w ight ≤ 700) organic molecules by light-directed solid-phase organic synthesis. The present invention is

- 22 -

particulary suited to the synthesis and assay of carbohydrate libraries, which can be extremely large due to the positional and stereochemical isomerism available to oligosaccharides (e.g., there are over 8.3 million tetrasaccharides derivable from the eight hexose monomers). Modifications and other applications of the present invention, which are apparent to those skilled in the fields of chemistry, biology, and the pharmaceutical sciences, are considered to be part of and within the scope of the present invention.

- 23 -

Claims

What I claim is:

1. A method for preparing a spatially-addressable chemical library array, comprising the steps of:

    a) providing a disc shaped planar substrate as a solid phase, wherein the substrate comprises reactive groups on its surface, the reactive groups being masked by laser-removable protecting groups, the surface further bearing formatting features that enable reproducible and accurate irradiation of surface sites by the laser of a recordable CD-ROM mechanism;

    b) preparing the library by light-directed synthesis on the planar substrate, wherein the light comprises laser radiation, and wherein the laser radiation is directed to pre-selected synthesis sites on the surface of the substrate by the CD-ROM mechanism under the control of a computer; and

    c) optionally removing some or all residual protecting groups.

2. A method for identifying the members of a spatially-addressable chemical library array that bind a given analyte, comprising the steps of:

    a) providing a spatially-addressable chemical library array bound to the surface of a disc-shaped substrate, wherein the surface bears formatting features that enable reproducible and accurate irradiation of synthesis sites by the laser of a CD-ROM mechanism, wherein the library members are attached to the synthesis sites, and wherein the synthesis sites are arranged on the surface so as to be individually irradiatable by the laser of the CD-ROM mechanism;

    b) exposing said substrate, bearing said library, to a solution of said analyte, under conditions that permit specific binding of said analyte to a subset of members of the library;

    c) removing the analyte that has not bound specifically to a subset of members of th library;

- 24 -

d)   optionally labeling the analyte that has bound to a subset of members of the library with a label, said label having the capacity to either emit light upon laser irradiation or to modify some measurable property of light from an interrogating laser;

e)   by means of a CD-ROM mechanism under computer control, irradiating a selected synthesis site on the surface of the substrate bearing the library and bearing the optionally labeled analyte bound to a subset of members of the library, with an interrogating laser;

f) by means of a CD-ROM mechanism under computer control, either detecting emitted light or measuring said measurable property in the reflected or transmitted beam of the interrogating laser, determining if the property has been changed due to the presence of analyte or label at the irradiated synthesis site, and recording the results of said determination;

g)   repeating steps (e) and (f) at other synthesis sites, irradiating synthesis sites systematically until all synthesis sites have been irradiated; and

h)   correlating the location of synthesis sites bearing analyte with the structure of the chemical library member synthesized at each such site, the structure being deduced by reference to a stored chronology.

3. A spatially addressable array of a library of chemical compounds on a planar substrate, said substrate being mountable in and carrying formatting features rendering it readable by a CD-ROM mechanism.

4. A method for preparing a spatially-addressable chemical library array, comprising the steps of:

a) providing a disc shaped planar substrate as a solid phase, wherein the substrate comprises reactive groups on its surface, the reactive groups being masked by laser-removable protecting groups, the planar substrate further bearing formatting features that nable reproducible and accurat irradiation of selected surface sites when the substrate is rotated;

- 25 -

b)   preparing the library by light-directed synthesis on the planar substrate, wherein the light comprises laser radiation, and wherein the laser radiation is directed to pre-selected synthesis sites on the surface of the substrate under the control of a computer; and

c)   optionally removing some or all residual protecting groups.

5.   A method for identifying the members of a spatially-addressable chemical library array that bind a given analyte, comprising the steps of:

a)   providing a spatially-addressable chemical library array bound to the surface of a disc-shaped substrate, wherein the surface bears formatting features that enable reproducible and selective laser irradiation of selected library members when the substrate is rotated, and wherein the library members are arranged on the surface so as to be individually irradiatable by the laser;

b)   exposing said substrate, bearing said library, to a solution of said analyte, under conditions that permit specific binding of said analyte to a subset of library members;

c)   removing the analyte that has not bound specifically to a subset of library members;

d)   optionally labeling the analyte that has bound to a subset of library members with a label, said label having the capacity to either emit light upon laser irradiatiation or to modify some measurable property of light from a laser;

e)   under computer control, and by reference to the formatting features, rotating the substrate and/or translating the laser radially, so as to bring a selected library member having a given spatial location and/or logical address on the substrate into a position where it may be irradiated by the laser;

f)   under computer control, irradiating said selected synth sis sit on the surface of the substrate with the laser;

g)   und r computer control, detecting emitted, reflected, or transmitted

- 26 -

light from the site of the selected library member;

h)   either measuring said measurable property in the reflected or transmitted light and determining if the property has been changed due to the presence of analyte or label at the site of the irradiated library member, or measuring the amount of emitted light;

i)   recording the spatial location and/or logical address of the library member, together with the result of the measurement of step (h) at the site of the selected library member, said result being indicative of the amount of analyte bound to the selected library member;

j)   repeating steps (e) through (i), irradiating library members systematically until all selected library members have been irradiated; and

k)   correlating the spatial location and/or logical address of library members to which the analyte has bound, with the structure of the library member at each such spatial location or logical address.

6.   A spatially addressable array of a library of chemical compounds on a disc shaped planar substrate, said substrate carrying formatting features that enable reproducible and selective laser irradiation of selected library members when the substrate is rotated.

7.   The method of claim 5, wherein the library members are oligonucleotides or oligonucleotide analogues, and wherein the analyte is DNA.

8.   The array of claim 6, wherein the chemical compounds are oligonucleotides or oligonucleotide analogues.